

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
7 February 2002 (07.02.2002)

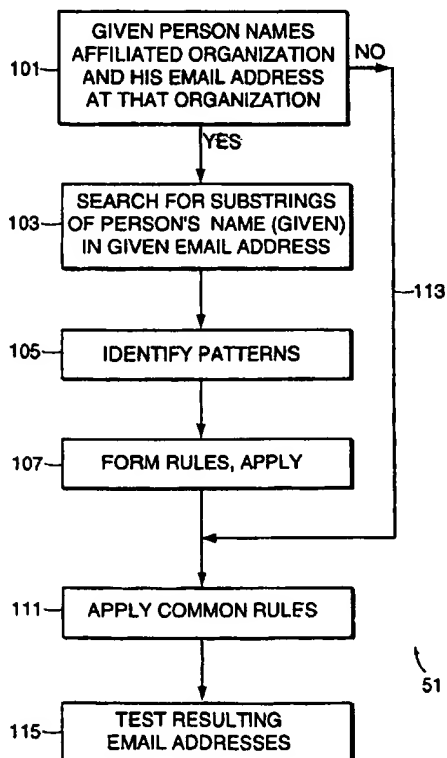
PCT

(10) International Publication Number
WO 02/010968 A3

- (51) International Patent Classification⁷: **G06F 17/30**, 17/60, H04L 12/58
- (21) International Application Number: **PCT/US01/41515**
- (22) International Filing Date: **31 July 2001 (31.07.2001)**
- (25) Filing Language: **English**
- (26) Publication Language: **English**
- (30) Priority Data:
- | | | |
|------------|------------------------------|----|
| 60/221,750 | 31 July 2000 (31.07.2000) | US |
| 09/704,080 | 1 November 2000 (01.11.2000) | US |
| 09/703,907 | 1 November 2000 (01.11.2000) | US |
| 09/768,869 | 24 January 2001 (24.01.2001) | US |
| 09/821,908 | 30 March 2001 (30.03.2001) | US |
| 09/918,312 | 20 July 2001 (20.07.2001) | US |
- (71) Applicant: **ELIYON TECHNOLOGIES CORPORATION** [US/US]; 810 Memorial Drive, Cambridge, MA 02139 (US).
- (72) Inventors: **STERN, Jonathan**; 9 Alton Place, Brookline, MA 02446 (US). **ROTHMAN-SHORE, Jeremy, W.**; 275 River Street, #2, Cambridge, MA 02139 (US). **KARADIMITRIOU, Kosmas**; 20 Brentwood Drive, Shrewsbury, MA 01545 (US). **DECARY, Michel**; 3807 Mackenzie, Montreal, Quebec H3S 1E4 (CA).
- (74) Agent: **WAKIMURA, Mary, Lou**; Hamilton, Brook, Smith & Reynolds, P.C., 530 Virginia Road, P.O. Box 9133, Concord, MA 01742-9133 (US).
- (81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC,

[Continued on next page]

(54) Title: **DATA MINING SYSTEM**



(57) Abstract: A computer automated method and system mines from a global computer network, information about people and organizations. The system includes automated crawling means, a distributor controlling the crawling means processing, an extractor storing extracted information of interest in a database, an integrator and post-processor. The integrator resolves duplicate information and combines related information in the database. The post-processor analyzes the database contents and generates (by deduction) missing information. Also disclosed is a method and means for generating a business email address of a person named in the database, from the respective organization named in the database for the subject person.

WO 02/010968 A3

AF3



LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW,
MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK,
SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.

Published:

— with international search report

(84) **Designated States (regional):** ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

(88) **Date of publication of the international search report:**
31 July 2003

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.